

Reliability and validity of NAPLAN (page 1 of 2)

The National Assessment Program – Literacy and Numeracy (NAPLAN) provides an assessment of the literacy and numeracy skills of students right across Australia. Processes have been put in place to ensure that NAPLAN is a valid and reliable measurement of students' literacy and numeracy ability. These processes are outlined below.

Purpose of NAPLAN

NAPLAN has a number of purposes including reporting national and jurisdictional achievements in literacy and numeracy as well as providing accurate and reliable measures of student and school performance. These purposes are carefully considered during the NAPLAN development process.

Development of NAPLAN tests

The tests have been developed using the nationally agreed Statements of Learning that reflect the core elements of the curriculum documents used in the different States and Territories. Specialist item writers are used by the NAPLAN test development contractors to develop test questions. When constructing test questions, writers must meet various specifications including curriculum coverage using National Statements of Learning and State/Territory curriculums as guidance, spread of item difficulties, length of tests, question types (multiple choice or open response) and the need to have test questions that are common to adjacent year levels.

Quality assurance

Proposed test questions are provided to all State and Territory government and non-government test authorities for review. Assessment and curriculum specialists and classroom teachers, as well as specialists in Indigenous education, English as a second language, disabilities and inclusive education, review the items to ensure that the test questions are accessible to all students. This process ensures that only those questions that meet stringent criteria go to trial.

NAPLAN trials

Samples of students from all jurisdictions participate in NAPLAN test trialling. The trial checks whether the items developed are suitable for inclusion in the NAPLAN tests, including whether there are sufficient items to cover the range of student abilities at each year level and the elements of each domain that need to be assessed. Data from the trialling are used, along with professional judgments from educational measurement, test construction and curriculum experts from all jurisdictions, to select the items for the final tests.

Expert advice

An Expert Advisory Group (EAG), consisting of five pre-eminent educational measurement experts, provides advice on test development and endorsement of the tests before they are finalised. The EAG also provides advice and endorsement of key technical methods and specifications for reporting, equating and standards.

Before the tests are signed off, the EAG is asked to review the tests thoroughly. The EAG may recommend changes prior to endorsement. The EAG provides a final report to ACARA that endorses the tests as satisfying the requirements.

Common scales

The use of common scales covering Years 3, 5, 7 and 9 for each area or domain assessed allows for an individual student's achievement to be mapped as he or she progresses through schooling. In order to enable the comparison between years of testing, the tests have to be able to be placed on the same common scales.

>> Continued on next page

Reliability and validity of NAPLAN (page 2 of 2)

Equating between test years

As no two tests can be exactly the same difficulty, a rigorous equating process is carried out so that the difficulty of the current year's tests can be adjusted, if need be, to the same level of difficulty as the previous year's tests.

In 2009, equating tests were developed so that future tests can be put on the same scale as previous years. The 2009 equating process used both on-shore and off-shore testing. Students in New Zealand sat the 2008 NAPLAN tests and the equating tests, while a sample of students in Australia who sat the 2009 NAPLAN tests also sat the equating tests. Using a combination of equating methods, the 2009 and equating tests were able to be placed on the same scales as the 2008 tests.

From 2010 a sample of students from each year, covering all States and Territories and school sectors, will sit the secure equating tests as well as the current year's tests. The equating tests will be delivered by specially trained independent test administrators. This ensures that the security of the equating tests can be preserved.

The equating process for NAPLAN was developed following advice from the EAG. Great care is taken to provide a high level of assurance as to the reliability of comparisons between years. The equating process enables us to say with confidence that any test difference has been taken into account before making statements about one year's results compared to the next.

Reliability

While no tests are 100 per cent reliable, there are routine methods for estimating the reliability of tests. These methods indicate that the reliability of NAPLAN tests is high and that they can be used with confidence and are fit for purpose.

The rigorous processes that are carried out during the development of NAPLAN each year ensure that the results are reliable and comparable between years. The results from NAPLAN provide a sound indication of the performance of Australian students in literacy and numeracy.

The Australian Council for Educational Research (ACER) and the EAG advise that NAPLAN represents 'world's best practice in the ability to measure student progress' (*Reporting and Comparing School Performance – ACER DEC2008*, available on the MCEECDYA website).